

## SEQUENCE LISTING

<110> University of Texas at San Antonio  
Louisiana State University Health Sciences Center  
Heidner, Hans Walter  
Klimstra, William Brown  
Ryman, Katherine Diana

<120> METHODS AND COMPOSITIONS COMPRISING PROTEIN L IMMUNOGLOBULIN  
BINDING DOMAINS FOR CELL-SPECIFIC TARGETING

<130> 9237.21WO

<160> 27

<170> PatentIn version 3.2

<210> 1  
<211> 2657  
<212> DNA  
<213> Peptostreptococcus magnus

<220>  
<221> CDS  
<222> (301)..(2460)  
<223> protein L

<220>  
<221> repeat\_unit  
<222> (592)..(819)  
<223> PpL1

<220>  
<221> repeat\_unit  
<222> (820)..(1035)  
<223> PpL2

<220>  
<221> repeat\_unit  
<222> (1036)..(1251)  
<223> PpL3

<220>  
<221> repeat\_unit  
<222> (1252)..(1467)  
<223> PpL4

<220>  
<221> repeat\_unit  
<222> (1468)..(1686)  
<223> PpL5

<400> 1  
atagtgtac aataaaggat ggcactgaac taccgaacct aacgcttgtg gacattgtct 60  
tggaaattt ggacagtgga cgaatcaaga acaccattaa ttaaattggc gaagttcgat 120

tgttgatcac	cctttttggg	taaacaataa	ctaaggaatg	gcaatatatt	tgcttgaaa	180										
cgaatttgat	ttaaatagca	ttgaatgcaa	aaaaatttaa	aaggaggaga	caaattccac	240										
ccttattaga	agggaaagttt	ccattgtcat	gatattatga	aaattaataa	gaaatttatta	300										
atg	gct	gca	ctt	gca	ggt	gca	att	gta	gta	aca	ggt	gga	gta	gga	tct	348
Met	Ala	Ala	Leu	Ala	Gly	Ala	Ile	Val	Val	Thr	Gly	Gly	Val	Gly	Ser	
1																15
10																
tac	gca	gct	gat	gaa	cct	att	gat	ctt	gaa	aaa	ctt	gaa	gaa	aaa	agg	396
Tyr	Ala	Ala	Asp	Glu	Pro	Ile	Asp	Leu	Glu	Lys	Leu	Glu	Glu	Lys	Arg	
20																30
25																
gat	aaa	gaa	aat	gta	gga	aat	tta	cca	aaa	ttc	gat	aat	gaa	gtt	aaa	444
Asp	Lys	Glu	Asn	Val	Gly	Asn	Leu	Pro	Lys	Phe	Asp	Asn	Glu	Val	Lys	
35																45
40																
gat	ggt	tca	gaa	aat	cca	atg	gct	aaa	tat	cca	gat	ttc	gat	gat	gaa	492
Asp	Gly	Ser	Glu	Asn	Pro	Met	Ala	Lys	Tyr	Pro	Asp	Phe	Asp	Asp	Glu	
50																55
55																60
gcc	agt	aca	aga	ttt	gaa	aca	gaa	aac	aat	gaa	ttt	gaa	gaa	aaa	aaa	540
Ala	Ser	Thr	Arg	Phe	Glu	Thr	Glu	Asn	Asn	Glu	Phe	Glu	Glu	Lys	Lys	
65																75
70																80
gtt	gtt	tct	gat	aac	ttt	ttt	gat	caa	tca	gaa	cat	ccg	ttt	gta	gaa	588
Val	Val	Ser	Asp	Asn	Phe	Phe	Asp	Gln	Ser	Glu	His	Pro	Phe	Val	Glu	
85																90
90																95
aat	aaa	gaa	aca	cca	gaa	aca	cca	gaa	act	gat	tca	gaa	gaa	gaa	gaa	636
Asn	Lys	Glu	Glu	Thr	Pro	Glu	Thr	Pro	Glu	Thr	Asp	Ser	Glu	Glu	Glu	
100																105
105																110
gta	aca	atc	aaa	gct	aac	cta	atc	ttt	gca	aat	gga	agc	aca	caa	act	684
Val	Thr	Ile	Lys	Ala	Asn	Leu	Ile	Phe	Ala	Asn	Gly	Ser	Thr	Gln	Thr	
115																120
120																125
gca	gaa	ttc	aaa	gga	aca	ttt	gaa	aaa	gca	aca	tca	gaa	gct	tat	gcg	732
Ala	Glu	Phe	Lys	Gly	Thr	Phe	Glu	Lys	Ala	Thr	Ser	Glu	Ala	Tyr	Ala	
130																135
135																140
tat	gca	gat	act	ttg	aag	aaa	gac	aat	gga	gaa	tat	act	gta	gat	gtt	780
Tyr	Ala	Asp	Thr	Leu	Lys	Lys	Asp	Asn	Gly	Glu	Tyr	Thr	Val	Asp	Val	
145																150
150																155
155																160
gca	gat	aaa	ggt	tat	act	tta	aat	att	aaa	ttt	gct	gga	aaa	gaa	aaa	828
Ala	Asp	Lys	Gly	Tyr	Thr	Leu	Asn	Ile	Lys	Phe	Ala	Gly	Lys	Glu	Lys	
165																170
170																175
aca	cca	gaa	cca	aaa	gaa	gaa	gtt	act	att	aaa	gca	aac	tta	atc	876	
Thr	Pro	Glu	Glu	Pro	Lys	Glu	Glu	Val	Thr	Ile	Lys	Ala	Asn	Leu	Ile	
180																185
185																190
tat	gca	gat	gga	aaa	aca	caa	aca	gca	gaa	ttc	aaa	gga	aca	ttt	gaa	924
Tyr	Ala	Asp	Gly	Lys	Thr	Gln	Thr	Ala	Glu	Phe	Lys	Gly	Thr	Phe	Glu	

195	200	205	
gaa gca aca gca gaa gca tac aga tat gca gat gca tta aag aag gac Glu Ala Thr Ala Glu Ala Tyr Arg Tyr Ala Asp Ala Leu Lys Lys Asp 210	215	220	972
aat gga gaa tat aca gta gac gtt gca gat aaa ggt tat act tta aat Asn Gly Glu Tyr Thr Val Asp Val Ala Asp Lys Gly Tyr Thr Leu Asn 225	230	235	1020
att aaa ttt gct gga aaa gaa aaa aca cca gaa gaa cca aaa gaa gaa Ile Lys Phe Ala Gly Lys Glu Lys Thr Pro Glu Glu Pro Lys Glu Glu 245	250	255	1068
gtt act att aaa gca aac tta atc tat gca gat gga aaa aca caa aca Val Thr Ile Lys Ala Asn Leu Ile Tyr Ala Asp Gly Lys Thr Gln Thr 260	265	270	1116
gca gaa ttc aaa gga aca ttt gaa gaa gca aca gca gaa gca tac aga Ala Glu Phe Lys Gly Thr Phe Glu Glu Ala Thr Ala Glu Ala Tyr Arg 275	280	285	1164
tat gct gac tta tta gca aaa gaa aat ggt aaa tat act gta gac gtt Tyr Ala Asp Leu Leu Ala Lys Glu Asn Gly Lys Tyr Thr Val Asp Val 290	295	300	1212
gca gat aaa ggt tat act tta aat att aaa ttt gct gga aaa gaa aaa Ala Asp Lys Gly Tyr Thr Leu Asn Ile Lys Phe Ala Gly Lys Glu Lys 305	310	315	1260
aca cca gaa gaa cca aaa gaa gaa gtt act att aaa gca aac tta atc Thr Pro Glu Glu Pro Lys Glu Glu Val Thr Ile Lys Ala Asn Leu Ile 325	330	335	1308
tat gca gat gga aaa act caa aca gca gag ttc aaa gga aca ttt gca Tyr Ala Asp Gly Lys Thr Gln Thr Ala Glu Phe Lys Gly Thr Phe Ala 340	345	350	1356
gaa gca aca gca gaa gca tac aga tac gct gac tta tta gca aaa gaa Glu Ala Thr Ala Glu Ala Tyr Arg Tyr Ala Asp Leu Leu Ala Lys Glu 355	360	365	1404
aat ggt aaa tat aca gca gac tta gaa gat ggt gga tac act att aat Asn Gly Lys Tyr Thr Ala Asp Leu Glu Asp Gly Gly Tyr Thr Ile Asn 370	375	380	1452
att aga ttt gca ggt aag aaa gtt gac gaa aaa cca gaa gaa aaa gaa Ile Arg Phe Ala Gly Lys Lys Val Asp Glu Lys Pro Glu Glu Lys Glu 385	390	395	1500
caa gta aca att aaa gaa aat ata tat ttt gaa gat gga aca gta caa Gln Val Thr Ile Lys Glu Asn Ile Tyr Phe Glu Asp Gly Thr Val Gln 405	410	415	1548
act gca aca ttt aaa gga aca ttt gca gaa gcg aca gca gaa gca tac Thr Ala Thr Phe Lys Gly Thr Phe Ala Glu Ala Thr Ala Glu Ala Tyr 420	425	430	1596

aga tat gca gat ttg tta tca aaa gaa cat ggt aaa tac aca gca gac	1644
Arg Tyr Ala Asp Leu Leu Ser Lys Glu His Gly Lys Tyr Thr Ala Asp	
435 440 445	
ttg gaa gat ggt gga tac act atc aac att aga ttt gct gga aaa gaa	1692
Leu Glu Asp Gly Gly Tyr Thr Ile Asn Ile Arg Phe Ala Gly Lys Glu	
450 455 460	
gaa cca gaa aca cca gaa aaa cca gaa gta caa gac gga tat gca	1740
Glu Pro Glu Glu Thr Pro Glu Lys Pro Glu Val Gln Asp Gly Tyr Ala	
465 470 475 480	
tca tac gaa gaa gct gaa gca gca gct aaa gaa gct ttg aaa aat gat	1788
Ser Tyr Glu Glu Ala Glu Ala Ala Lys Glu Ala Leu Lys Asn Asp	
485 490 495	
gat gta aat aaa tca tat act att aga caa ggt gca gat gga aga tat	1836
Asp Val Asn Lys Ser Tyr Thr Ile Arg Gln Gly Ala Asp Gly Arg Tyr	
500 505 510	
tac tat gta tta tca cca gta gaa gct gaa gaa gaa aaa cca gaa gca	1884
Tyr Tyr Val Leu Ser Pro Val Glu Ala Glu Glu Lys Pro Glu Ala	
515 520 525	
caa aat gga tat gca aca tac gaa gaa gca gaa gca gca gct aag aaa	1932
Gln Asn Gly Tyr Ala Thr Tyr Glu Glu Ala Glu Ala Ala Lys Lys	
530 535 540	
gct ttg gaa aat gat cca atc aat aaa tct tac agc ata aga caa ggt	1980
Ala Leu Glu Asn Asp Pro Ile Asn Lys Ser Tyr Ser Ile Arg Gln Gly	
545 550 555 560	
gca gat gga aga tac tac tat gta tta tca cca gta gaa gct gaa act	2028
Ala Asp Gly Arg Tyr Tyr Val Leu Ser Pro Val Glu Ala Glu Thr	
565 570 575	
cct gaa aaa cca gta gaa cca tca gaa cca tca aca cca gat gta cca	2076
Pro Glu Lys Pro Val Glu Pro Ser Glu Pro Ser Thr Pro Asp Val Pro	
580 585 590	
tca aac cca tca aat cca tca aca cca gat gtt cca tca act cct gat	2124
Ser Asn Pro Ser Asn Pro Ser Thr Pro Asp Val Pro Ser Thr Pro Asp	
595 600 605	
gta cca tca aat cct tca act cca gaa gtt cct tca aac cca tca act	2172
Val Pro Ser Asn Pro Ser Thr Pro Glu Val Pro Ser Asn Pro Ser Thr	
610 615 620	
cct ggt aat gaa gaa aaa cct ggt aac gaa caa aaa cct ggt aac gaa	2220
Pro Gly Asn Glu Glu Lys Pro Gly Asn Glu Gln Lys Pro Gly Asn Glu	
625 630 635 640	
caa aaa cct ggt aac gaa caa aaa cct ggt aat gaa caa aaa cct ggt	2268
Gln Lys Pro Gly Asn Glu Gln Lys Pro Gly Asn Glu Gln Lys Pro Gly	
645 650 655	

aat gaa caa aaa cca gac caa cct tca aaa cca gaa aaa gaa gaa aat	2316	
Asn Glu Gln Lys Pro Asp Gln Pro Ser Lys Pro Glu Lys Glu Glu Asn		
660	665	670
ggt aag ggt gga gta gat tct cca aag aaa aaa gaa aaa gct gca tta	2364	
Gly Lys Gly Gly Val Asp Ser Pro Lys Lys Glu Lys Ala Ala Leu		
675	680	685
cca aaa gct ggt agc gaa gct gaa atc tta aca tta gca gca gct tca	2412	
Pro Lys Ala Gly Ser Glu Ala Glu Ile Leu Thr Leu Ala Ala Ala Ser		
690	695	700
tta tca agc gtt gca ggt gct ttc att tca ctt aaa aaa cgt aaa taa	2460	
Leu Ser Ser Val Ala Gly Ala Phe Ile Ser Leu Lys Lys Arg Lys		
705	710	715
ttaatcttag ataaaaaaaata gattaatcta aaaaaatggg acttttataa gtcccatttt	2520	
tgattgcaat gaaactgata caaaaaatgt atcagtttt tcatttacgc ttattttcct	2580	
gtgagtatgt ccaagttgt cgtataatca tctatcaaaa gcctggcgag tttcgctttt	2640	
tggattttgt cttcgaa	2657	
<210> 2		
<211> 719		
<212> PRT		
<213> Peptostreptococcus magnus		
<400> 2		
Met Ala Ala Leu Ala Gly Ala Ile Val Val Thr Gly Gly Val Gly Ser		
1	5	10
15		
Tyr Ala Ala Asp Glu Pro Ile Asp Leu Glu Lys Leu Glu Glu Lys Arg		
20	25	30
Asp Lys Glu Asn Val Gly Asn Leu Pro Lys Phe Asp Asn Glu Val Lys		
35	40	45
Asp Gly Ser Glu Asn Pro Met Ala Lys Tyr Pro Asp Phe Asp Asp Glu		
50	55	60
Ala Ser Thr Arg Phe Glu Thr Glu Asn Asn Glu Phe Glu Glu Lys Lys		
65	70	75
80		
Val Val Ser Asp Asn Phe Phe Asp Gln Ser Glu His Pro Phe Val Glu		
85	90	95
Asn Lys Glu Glu Thr Pro Glu Thr Pro Glu Thr Asp Ser Glu Glu Glu		

100

105

110

Val Thr Ile Lys Ala Asn Leu Ile Phe Ala Asn Gly Ser Thr Gln Thr  
115 120 125

Ala Glu Phe Lys Gly Thr Phe Glu Lys Ala Thr Ser Glu Ala Tyr Ala  
130 135 140

Tyr Ala Asp Thr Leu Lys Lys Asp Asn Gly Glu Tyr Thr Val Asp Val  
145 150 155 160

Ala Asp Lys Gly Tyr Thr Leu Asn Ile Lys Phe Ala Gly Lys Glu Lys  
165 170 175

Thr Pro Glu Glu Pro Lys Glu Glu Val Thr Ile Lys Ala Asn Leu Ile  
180 185 190

Tyr Ala Asp Gly Lys Thr Gln Thr Ala Glu Phe Lys Gly Thr Phe Glu  
195 200 205

Glu Ala Thr Ala Glu Ala Tyr Arg Tyr Ala Asp Ala Leu Lys Lys Asp  
210 215 220

Asn Gly Glu Tyr Thr Val Asp Val Ala Asp Lys Gly Tyr Thr Leu Asn  
225 230 235 240

Ile Lys Phe Ala Gly Lys Glu Lys Thr Pro Glu Glu Pro Lys Glu Glu  
245 250 255

Val Thr Ile Lys Ala Asn Leu Ile Tyr Ala Asp Gly Lys Thr Gln Thr  
260 265 270

Ala Glu Phe Lys Gly Thr Phe Glu Glu Ala Thr Ala Glu Ala Tyr Arg  
275 280 285

Tyr Ala Asp Leu Leu Ala Lys Glu Asn Gly Lys Tyr Thr Val Asp Val  
290 295 300

Ala Asp Lys Gly Tyr Thr Leu Asn Ile Lys Phe Ala Gly Lys Glu Lys  
305 310 315 320

Thr Pro Glu Glu Pro Lys Glu Glu Val Thr Ile Lys Ala Asn Leu Ile  
325 330 335

Tyr Ala Asp Gly Lys Thr Gln Thr Ala Glu Phe Lys Gly Thr Phe Ala  
340 345 350

Glu Ala Thr Ala Glu Ala Tyr Arg Tyr Ala Asp Leu Leu Ala Lys Glu  
355 360 365

Asn Gly Lys Tyr Thr Ala Asp Leu Glu Asp Gly Gly Tyr Thr Ile Asn  
370 375 380

Ile Arg Phe Ala Gly Lys Lys Val Asp Glu Lys Pro Glu Glu Lys Glu  
385 390 395 400

Gln Val Thr Ile Lys Glu Asn Ile Tyr Phe Glu Asp Gly Thr Val Gln  
405 410 415

Thr Ala Thr Phe Lys Gly Thr Phe Ala Glu Ala Thr Ala Glu Ala Tyr  
420 425 430

Arg Tyr Ala Asp Leu Leu Ser Lys Glu His Gly Lys Tyr Thr Ala Asp  
435 440 445

Leu Glu Asp Gly Gly Tyr Thr Ile Asn Ile Arg Phe Ala Gly Lys Glu  
450 455 460

Glu Pro Glu Glu Thr Pro Glu Lys Pro Glu Val Gln Asp Gly Tyr Ala  
465 470 475 480

Ser Tyr Glu Glu Ala Glu Ala Ala Lys Glu Ala Leu Lys Asn Asp  
485 490 495

Asp Val Asn Lys Ser Tyr Thr Ile Arg Gln Gly Ala Asp Gly Arg Tyr  
500 505 510

Tyr Tyr Val Leu Ser Pro Val Glu Ala Glu Glu Lys Pro Glu Ala  
515 520 525

Gln Asn Gly Tyr Ala Thr Tyr Glu Glu Ala Glu Ala Ala Lys Lys  
530 535 540

Ala Leu Glu Asn Asp Pro Ile Asn Lys Ser Tyr Ser Ile Arg Gln Gly  
545 550 555 560

Ala Asp Gly Arg Tyr Tyr Tyr Val Leu Ser Pro Val Glu Ala Glu Thr  
 565 570 575

Pro Glu Lys Pro Val Glu Pro Ser Glu Pro Ser Thr Pro Asp Val Pro  
 580 585 590

Ser Asn Pro Ser Asn Pro Ser Thr Pro Asp Val Pro Ser Thr Pro Asp  
 595 600 605

Val Pro Ser Asn Pro Ser Thr Pro Glu Val Pro Ser Asn Pro Ser Thr  
 610 615 620

Pro Gly Asn Glu Glu Lys Pro Gly Asn Glu Gln Lys Pro Gly Asn Glu  
 625 630 635 640

Gln Lys Pro Gly Asn Glu Gln Lys Pro Gly Asn Glu Gln Lys Pro Gly  
 645 650 655

Asn Glu Gln Lys Pro Asp Gln Pro Ser Lys Pro Glu Lys Glu Glu Asn  
 660 665 670

Gly Lys Gly Gly Val Asp Ser Pro Lys Lys Lys Glu Lys Ala Ala Leu  
 675 680 685

Pro Lys Ala Gly Ser Glu Ala Glu Ile Leu Thr Leu Ala Ala Ala Ser  
 690 695 700

Leu Ser Ser Val Ala Gly Ala Phe Ile Ser Leu Lys Lys Arg Lys  
 705 710 715

<210> 3  
 <211> 228  
 <212> DNA  
 <213> *Peptostreptococcus magnus*

<220>  
 <221> CDS  
 <222> (1)..(228)  
 <223> PpL1

<400> 3  
 aaa gaa gaa aca cca gaa aca cca gaa act gat tca gaa gaa gaa gta  
 Lys Glu Glu Thr Pro Glu Thr Pro Glu Thr Asp Ser Glu Glu Glu Val  
 1 5 10 15

aca atc aaa gct aac cta atc ttt gca aat gga agc aca caa act gca 96  
 Thr Ile Lys Ala Asn Leu Ile Phe Ala Asn Gly Ser Thr Gln Thr Ala  
 20 25 30

gaa ttc aaa gga aca ttt gaa aaa gca aca tca gaa gct tat gcg tat 144  
 Glu Phe Lys Gly Thr Phe Glu Lys Ala Thr Ser Glu Ala Tyr Ala Tyr  
 35 40 45

gca gat act ttg aag aaa gac aat gga gaa tat act gta gat gtt gca 192  
 Ala Asp Thr Leu Lys Lys Asp Asn Gly Glu Tyr Thr Val Asp Val Ala  
 50 55 60

gat aaa ggt tat act tta aat att aaa ttt gct gga 228  
 Asp Lys Gly Tyr Thr Leu Asn Ile Lys Phe Ala Gly  
 65 70 75

<210> 4  
<211> 76  
<212> PRT  
<213> Peptostreptococcus magnus

<400> 4

Lys Glu Glu Thr Pro Glu Thr Pro Glu Thr Asp Ser Glu Glu Glu Val  
 1 5 10 15

Thr Ile Lys Ala Asn Leu Ile Phe Ala Asn Gly Ser Thr Gln Thr Ala  
 20 25 30

Glu Phe Lys Gly Thr Phe Glu Lys Ala Thr Ser Glu Ala Tyr Ala Tyr  
 35 40 45

Ala Asp Thr Leu Lys Lys Asp Asn Gly Glu Tyr Thr Val Asp Val Ala  
 50 55 60

Asp Lys Gly Tyr Thr Leu Asn Ile Lys Phe Ala Gly  
 65 70 75

<210> 5  
<211> 216  
<212> DNA  
<213> Peptostreptococcus magnus

<220>  
<221> CDS  
<222> (1)..(216)  
<223> PpL2

<400> 5  
aaa gaa aaa aca cca gaa gaa cca aaa gaa gaa gtt act att aaa gca 48

Lys	Glu	Lys	Thr	Pro	Glu	Glu	Pro	Lys	Glu	Glu	Val	Thr	Ile	Lys	Ala	
1				5				10					15			
aac tta atc tat gca gat gga aaa aca caa aca gca gaa ttc aaa gga															96	
Asn	Leu	Ile	Tyr	Ala	Asp	Gly	Lys	Thr	Gln	Thr	Ala	Glu	Phe	Lys	Gly	
							20			25			30			
aca ttt gaa gaa gca aca gca gaa gca tac aga tat gca gat gca tta															144	
Thr	Phe	Glu	Glu	Ala	Thr	Ala	Glu	Ala	Tyr	Arg	Tyr	Ala	Asp	Ala	Leu	
							35			40		45				
aag aag gac aat gga gaa tat aca gta gac gtt gca gat aaa ggt tat															192	
Lys	Lys	Asp	Asn	Gly	Glu	Tyr	Thr	Val	Asp	Val	Ala	Asp	Lys	Gly	Tyr	
						50			55		60					
act tta aat att aaa ttt gct gga															216	
Thr	Leu	Asn	Ile	Lys	Phe	Ala	Gly									
						65			70							
<210> 6																
<211> 72																
<212> PRT																
<213> Peptostreptococcus magnus																
<400> 6																
Lys	Glu	Lys	Thr	Pro	Glu	Glu	Pro	Lys	Glu	Glu	Val	Thr	Ile	Lys	Ala	
1				5				10					15			
Asn Leu Ile Tyr Ala Asp Gly Lys Thr Gln Thr Ala Glu Phe Lys Gly																
				20				25			30					
Thr Phe Glu Glu Ala Thr Ala Glu Ala Tyr Arg Tyr Ala Asp Ala Leu																
				35				40			45					
Lys Lys Asp Asn Gly Glu Tyr Thr Val Asp Val Ala Asp Lys Gly Tyr																
				50				55			60					
Thr Leu Asn Ile Lys Phe Ala Gly																
				65				70								
<210> 7																
<211> 216																
<212> DNA																
<213> Peptostreptococcus magnus																
<220>																
<221> CDS																
<222> (1) .. (216)																
<223> PpL3																

<400> 7  
 aaa gaa aaa aca cca gaa gaa cca aaa gaa gaa gtt act att aaa gca 48  
 Lys Glu Lys Thr Pro Glu Glu Pro Lys Glu Glu Val Thr Ile Lys Ala  
 1 5 10 15

aac tta atc tat gca gat gga aaa aca caa aca gca gaa ttc aaa gga 96  
 Asn Leu Ile Tyr Ala Asp Gly Lys Thr Gln Thr Ala Glu Phe Lys Gly  
 20 25 30

aca ttt gaa gaa gca aca gca gaa gca tac aga tat gct gac tta tta 144  
 Thr Phe Glu Glu Ala Thr Ala Glu Ala Tyr Arg Tyr Ala Asp Leu Leu  
 35 40 45

gca aaa gaa aat ggt aaa tat aca gta gac gtt gca gat aaa ggt tat 192  
 Ala Lys Glu Asn Gly Lys Tyr Thr Val Asp Val Ala Asp Lys Gly Tyr  
 50 55 60

act tta aat att aaa ttt gct gga 216  
 Thr Leu Asn Ile Lys Phe Ala Gly  
 65 70

<210> 8  
 <211> 72  
 <212> PRT  
 <213> Peptostreptococcus magnus

<400> 8  
 Lys Glu Lys Thr Pro Glu Glu Pro Lys Glu Glu Val Thr Ile Lys Ala 15  
 1 5 10

Asn Leu Ile Tyr Ala Asp Gly Lys Thr Gln Thr Ala Glu Phe Lys Gly  
 20 25 30

Thr Phe Glu Glu Ala Thr Ala Glu Ala Tyr Arg Tyr Ala Asp Leu Leu  
 35 40 45

Ala Lys Glu Asn Gly Lys Tyr Thr Val Asp Val Ala Asp Lys Gly Tyr  
 50 55 60

Thr Leu Asn Ile Lys Phe Ala Gly  
 65 70

<210> 9  
 <211> 216  
 <212> DNA  
 <213> Peptostreptococcus magnus

<220>

<221> CDS  
 <222> (1)..(216)  
 <223> PpL4

<400> 9  
 aaa gaa aaa aca cca gaa gaa cca aaa gaa gaa gtt act att aaa gca 48  
 Lys Glu Lys Thr Pro Glu Glu Pro Lys Glu Glu Val Thr Ile Lys Ala  
 1 5 10 15  
 aac tta atc tat gca gat gga aaa act caa aca gca gag ttc aaa gga 96  
 Asn Leu Ile Tyr Ala Asp Gly Lys Thr Gln Thr Ala Glu Phe Lys Gly  
 20 25 30  
 aca ttt gca gaa gca aca gca gaa gca tac aga tac gct gac tta tta 144  
 Thr Phe Ala Glu Ala Thr Ala Glu Ala Tyr Arg Tyr Ala Asp Leu Leu  
 35 40 45  
 gca aaa gaa aat ggt aaa tat aca gca gac tta gaa gat ggt gga tac 192  
 Ala Lys Glu Asn Gly Lys Tyr Thr Ala Asp Leu Glu Asp Gly Gly Tyr  
 50 55 60  
 act att aat att aga ttt gca ggt 216  
 Thr Ile Asn Ile Arg Phe Ala Gly  
 65 70

<210> 10  
 <211> 72  
 <212> PRT  
 <213> Peptostreptococcus magnus

<400> 10  
 Lys Glu Lys Thr Pro Glu Glu Pro Lys Glu Glu Val Thr Ile Lys Ala  
 1 5 10 15  
 Asn Leu Ile Tyr Ala Asp Gly Lys Thr Gln Thr Ala Glu Phe Lys Gly  
 20 25 30  
 Thr Phe Ala Glu Ala Thr Ala Glu Ala Tyr Arg Tyr Ala Asp Leu Leu  
 35 40 45  
 Ala Lys Glu Asn Gly Lys Tyr Thr Ala Asp Leu Glu Asp Gly Gly Tyr  
 50 55 60  
 Thr Ile Asn Ile Arg Phe Ala Gly  
 65 70

<210> 11  
 <211> 219  
 <212> DNA  
 <213> Peptostreptococcus magnus

<220>  
 <221> CDS  
 <222> (1)..(219)  
 <223> PpL5

<400> 11  
 aag aaa gtt gac gaa aaa cca gaa gaa aaa gaa caa gta aca att aaa 48  
 Lys Lys Val Asp Glu Lys Pro Glu Glu Lys Glu Gln Val Thr Ile Lys  
 1 5 10 15  
  
 gaa aat ata tat ttt gaa gat gga aca gta caa act gca aca ttt aaa 96  
 Glu Asn Ile Tyr Phe Glu Asp Gly Thr Val Gln Thr Ala Thr Phe Lys  
 20 25 30  
  
 gga aca ttt gca gaa gcg aca gca gaa gca tac aga tat gca gat ttg 144  
 Gly Thr Phe Ala Glu Ala Thr Ala Glu Ala Tyr Arg Tyr Ala Asp Leu  
 35 40 45  
  
 tta tca aaa gaa cat ggt aaa tac aca gca gac ttg gaa gat ggt gga 192  
 Leu Ser Lys Glu His Gly Lys Tyr Thr Ala Asp Leu Glu Asp Gly Gly  
 50 55 60  
  
 tac act atc aac att aga ttt gct gga 219  
 Tyr Thr Ile Asn Ile Arg Phe Ala Gly  
 65 70

<210> 12  
 <211> 73  
 <212> PRT  
 <213> Peptostreptococcus magnus

<400> 12  
  
 Lys Lys Val Asp Glu Lys Pro Glu Glu Lys Glu Gln Val Thr Ile Lys  
 1 5 10 15  
  
 Glu Asn Ile Tyr Phe Glu Asp Gly Thr Val Gln Thr Ala Thr Phe Lys  
 20 25 30  
  
 Gly Thr Phe Ala Glu Ala Thr Ala Glu Ala Tyr Arg Tyr Ala Asp Leu  
 35 40 45  
  
 Leu Ser Lys Glu His Gly Lys Tyr Thr Ala Asp Leu Glu Asp Gly Gly  
 50 55 60  
  
 Tyr Thr Ile Asn Ile Arg Phe Ala Gly  
 65 70  
  
 <210> 13

```

<211> 1269
<212> DNA
<213> Sindbis virus

<220>
<221> misc_feature
<222> (1)..(1269)
<223> E2 potein gene from strain AR339

<220>
<221> CDS
<222> (1)..(1269)
<223> E2 glycoprotein

<220>
<221> variation
<222> (3)..(3)
<223> a in AR339; c in HRsp

<220>
<221> variation
<222> (8)..(8)
<223> c in AR339; t in HRsp

<220>
<221> variation
<222> (68)..(68)
<223> a in AR339; t in HRsp

<220>
<221> variation
<222> (208)..(208)
<223> g in AR339; a in HRsp

<220>
<221> variation
<222> (514)..(514)
<223> g in AR339; a in HRsp

<400> 13
aga gtc act gac gac ttt acc ctg acc agc ccc tac ttg ggc aca tgc      48
Arg Val Thr Asp Asp Phe Thr Leu Thr Ser Pro Tyr Leu Gly Thr Cys
1          5          10          15

tcg tac tgc cac cat act gaa ccg tgc ttc agc cct gtt aag atc gag      96
Ser Tyr Cys His His Thr Glu Pro Cys Phe Ser Pro Val Lys Ile Glu
20         25         30

cag gtc tgg gac gaa gcg gac gat aac acc ata cgc ata cag act tcc      144
Gln Val Trp Asp Glu Ala Asp Asp Asn Thr Ile Arg Ile Gln Thr Ser
35         40         45

gcc cag ttt gga tac gac caa agc gga gca gca agc gca aac aag tac      192
Ala Gln Phe Gly Tyr Asp Gln Ser Gly Ala Ala Ser Ala Asn Lys Tyr
50         55         60

```

cgc tac atg tcg ctt gag cag gat cac acc gtt aaa gaa ggc acc atg	240
Arg Tyr Met Ser Leu Glu Gln Asp His Thr Val Lys Glu Gly Thr Met	
65 70 75 80	
gat gac atc aag att agc acc tca gga ccg tgt aga agg ctt agc tac	288
Asp Asp Ile Lys Ile Ser Thr Ser Gly Pro Cys Arg Arg Leu Ser Tyr	
85 90 95	
aaa gga tac ttt ctc ctc gca aaa tgc cct cca ggg gac agc gta acg	336
Lys Gly Tyr Phe Leu Leu Ala Lys Cys Pro Pro Gly Asp Ser Val Thr	
100 105 110	
gtt agc ata gtg agt agc aac tca gca acg tca tgt aca ctg gcc cgc	384
Val Ser Ile Val Ser Ser Asn Ser Ala Thr Ser Cys Thr Leu Ala Arg	
115 120 125	
aag ata aaa cca aaa ttc gtg gga cgg gaa aaa tat gat cta cct ccc	432
Lys Ile Lys Pro Lys Phe Val Gly Arg Glu Lys Tyr Asp Leu Pro Pro	
130 135 140	
gtt cac ggt aaa aaa att cct tgc aca gtg tac gac cgt ctg aaa gaa	480
Val His Gly Lys Ile Pro Cys Thr Val Tyr Asp Arg Leu Lys Glu	
145 150 155 160	
aca act gca ggc tac atc act atg cac agg ccg gga ccg cac gct tat	528
Thr Thr Ala Gly Tyr Ile Thr Met His Arg Pro Gly Pro His Ala Tyr	
165 170 175	
aca tcc tac ctg gaa gaa tca tca ggg aaa gtt tac gca aag ccg cca	576
Thr Ser Tyr Leu Glu Glu Ser Ser Gly Lys Val Tyr Ala Lys Pro Pro	
180 185 190	
tct ggg aag aac att acg tat gag tgc aag tgc ggc gac tac aag acc	624
Ser Gly Lys Asn Ile Thr Tyr Glu Cys Lys Cys Gly Asp Tyr Lys Thr	
195 200 205	
gga acc gtt tcg acc cgc acc gaa atc act ggt tgc acc gcc atc aag	672
Gly Thr Val Ser Thr Arg Thr Glu Ile Thr Gly Cys Thr Ala Ile Lys	
210 215 220	
cag tgc gtc gcc tat aag agc gac caa acg aag tgg gtc ttc aac tca	720
Gln Cys Val Ala Tyr Lys Ser Asp Gln Thr Lys Trp Val Phe Asn Ser	
225 230 235 240	
ccg gac ttg atc aga cat gac gac cac acg gcc caa ggg aaa ttg cat	768
Pro Asp Leu Ile Arg His Asp Asp His Thr Ala Gln Gly Lys Leu His	
245 250 255	
ttg cct ttc aag ttg atc ccg agt acc tgc atg gtc cct gtt gcc cac	816
Leu Pro Phe Lys Leu Ile Pro Ser Thr Cys Met Val Pro Val Ala His	
260 265 270	
gcg ccg aat gta ata cat ggc ttt aaa cac atc agc ctc caa tta gat	864
Ala Pro Asn Val Ile His Gly Phe Lys His Ile Ser Leu Gln Leu Asp	
275 280 285	
aca gac cac ttg aca ttg ctc acc acc agg aga cta ggg gca aac ccg	912

Thr	Asp	His	Leu	Thr	Leu	Leu	Thr	Thr	Arg	Arg	Leu	Gly	Ala	Asn	Pro	
290				295					300							
gaa	cca	acc	act	gaa	tgg	atc	gtc	gga	aag	acg	gtc	aga	aac	ttc	acc	960
Glu	Pro	Thr	Thr	Glu	Trp	Ile	Val	Gly	Lys	Thr	Val	Arg	Asn	Phe	Thr	
305				310				315				320				
gtc	gac	cga	gat	ggc	ctg	gaa	tac	ata	tgg	gga	aat	cat	gag	cca	gtg	1008
Val	Asp	Arg	Asp	Gly	Leu	Glu	Tyr	Ile	Trp	Gly	Asn	His	Glu	Pro	Val	
								325		330				335		
agg	gtc	tat	gcc	caa	gag	tca	gca	cca	gga	gac	cct	cac	gga	tgg	cca	1056
Arg	Val	Tyr	Ala	Gln	Glu	Ser	Ala	Pro	Gly	Asp	Pro	His	Gly	Trp	Pro	
								340		345				350		
cac	gaa	ata	gta	cag	cat	tac	cat	cgc	cat	cct	gtg	tac	acc	atc		1104
His	Glu	Ile	Val	Gln	His	Tyr	Tyr	His	Arg	His	Pro	Val	Tyr	Thr	Ile	
								355		360				365		
tta	gcc	gtc	gca	tca	gct	acc	gtg	gcg	atg	atg	att	ggc	gta	acc	gtt	1152
Leu	Ala	Val	Ala	Ser	Ala	Thr	Val	Ala	Met	Met	Ile	Gly	Val	Thr	Val	
								370		375				380		
gca	gtg	tta	tgt	gcc	tgt	aaa	gcg	cgc	cgt	gag	tgc	ctg	acg	cca	tac	1200
Ala	Val	Leu	Cys	Ala	Cys	Lys	Ala	Arg	Arg	Glu	Cys	Leu	Thr	Pro	Tyr	
								385		390				400		
gcc	ctg	gcc	cca	aac	gcc	gta	atc	cca	act	tgc	ctg	gca	ctc	ttg	tgc	1248
Ala	Leu	Ala	Pro	Asn	Ala	Val	Ile	Pro	Thr	Ser	Leu	Ala	Leu	Leu	Cys	
								405		410				415		
tgc	gtt	agg	tcg	gcc	aat	gct										1269
Cys	Val	Arg	Ser	Ala	Asn	Ala										
								420								

<210> 14  
 <211> 423  
 <212> PRT  
 <213> Sindbis virus

<400> 14

Arg	Val	Thr	Asp	Asp	Phe	Thr	Leu	Thr	Ser	Pro	Tyr	Leu	Gly	Thr	Cys
1					5			10				15			

Ser	Tyr	Cys	His	His	Thr	Glu	Pro	Cys	Phe	Ser	Pro	Val	Lys	Ile	Glu
					20			25				30			

Gln	Val	Trp	Asp	Glu	Ala	Asp	Asp	Asn	Thr	Ile	Arg	Ile	Gln	Thr	Ser
								35		40		45			

Ala	Gln	Phe	Gly	Tyr	Asp	Gln	Ser	Gly	Ala	Ala	Ser	Ala	Asn	Lys	Tyr
					50		55				60				

Arg Tyr Met Ser Leu Glu Gln Asp His Thr Val Lys Glu Gly Thr Met  
65 70 75 80

Asp Asp Ile Lys Ile Ser Thr Ser Gly Pro Cys Arg Arg Leu Ser Tyr  
85 90 95

Lys Gly Tyr Phe Leu Leu Ala Lys Cys Pro Pro Gly Asp Ser Val Thr  
100 105 110

Val Ser Ile Val Ser Ser Asn Ser Ala Thr Ser Cys Thr Leu Ala Arg  
115 120 125

Lys Ile Lys Pro Lys Phe Val Gly Arg Glu Lys Tyr Asp Leu Pro Pro  
130 135 140

Val His Gly Lys Lys Ile Pro Cys Thr Val Tyr Asp Arg Leu Lys Glu  
145 150 155 160

Thr Thr Ala Gly Tyr Ile Thr Met His Arg Pro Gly Pro His Ala Tyr  
165 170 175

Thr Ser Tyr Leu Glu Glu Ser Ser Gly Lys Val Tyr Ala Lys Pro Pro  
180 185 190

Ser Gly Lys Asn Ile Thr Tyr Glu Cys Lys Cys Gly Asp Tyr Lys Thr  
195 200 205

Gly Thr Val Ser Thr Arg Thr Glu Ile Thr Gly Cys Thr Ala Ile Lys  
210 215 220

Gln Cys Val Ala Tyr Lys Ser Asp Gln Thr Lys Trp Val Phe Asn Ser  
225 230 235 240

Pro Asp Leu Ile Arg His Asp Asp His Thr Ala Gln Gly Lys Leu His  
245 250 255

Leu Pro Phe Lys Leu Ile Pro Ser Thr Cys Met Val Pro Val Ala His  
260 265 270

Ala Pro Asn Val Ile His Gly Phe Lys His Ile Ser Leu Gln Leu Asp  
275 280 285

Thr Asp His Leu Thr Leu Leu Thr Thr Arg Arg Leu Gly Ala Asn Pro  
290 295 300

Glu Pro Thr Thr Glu Trp Ile Val Gly Lys Thr Val Arg Asn Phe Thr  
305 310 315 320

Val Asp Arg Asp Gly Leu Glu Tyr Ile Trp Gly Asn His Glu Pro Val  
325 330 335

Arg Val Tyr Ala Gln Glu Ser Ala Pro Gly Asp Pro His Gly Trp Pro  
340 345 350

His Glu Ile Val Gln His Tyr Tyr His Arg His Pro Val Tyr Thr Ile  
355 360 365

Leu Ala Val Ala Ser Ala Thr Val Ala Met Met Ile Gly Val Thr Val  
370 375 380

Ala Val Leu Cys Ala Cys Lys Ala Arg Arg Glu Cys Leu Thr Pro Tyr  
385 390 395 400

Ala Leu Ala Pro Asn Ala Val Ile Pro Thr Ser Leu Ala Leu Leu Cys  
405 410 415

Cys Val Arg Ser Ala Asn Ala  
420

<210> 15  
<211> 11703  
<212> DNA  
<213> Sindbis virus

<220>  
<221> misc\_feature  
<222> (1)..(11703)  
<223> Sindbis virus HRsp and wild-type strains complete genome

<220>  
<221> CDS  
<222> (8439)..(8630)  
<223> E3 protein

<220>  
<221> CDS  
<222> (8631)..(9899)  
<223> E2 protein

<400> 15  
 attgacggcg tagtacacac tattgaatca aacagccgac caattgcact accatcacaa 60  
 tggagaagcc agtagtaaac gtagacgtag acccccagag tccgtttgcgt gtgcaactgc 120  
 aaaaaagctt cccgcaattt gaggttagtag cacagcaggt cactccaaat gaccatgcta 180  
 atgccagagc attttcgcat ctggccagta aactaatcga gctggagggtt cctaccacag 240  
 cgacgatctt ggacataggc agcgcaccgg ctcgtagaat gttttccgag caccagtatc 300  
 attgtgtctg ccccatgcgt agtccagaag acccggaccg catgatgaaa tacgccagta 360  
 aactggcggaa aaaagcgtgc aagattacaa acaagaactt gcatgagaag attaaggatc 420  
 tccggaccgt acttgatacg ccggatgctg aaacaccatc gctctgcttt cacaacgatg 480  
 ttacctgcaa catgcgtgcc gaatattccg tcatgcagga cgtgtatatc aacgctcccg 540  
 gaactatcta tcatcaggct atgaaaggcg tgcggaccct gtactggatt ggcttcgaca 600  
 ccacccagtt catgttctcg gctatggcag gttcgtaccc tgcgtacaac accaactggg 660  
 ccgacgagaa agtccttgaa gcgcgtaaca tcggactttg cagcacaaag ctgagtgaag 720  
 gtaggacagg aaaattgtcg ataatgagga agaaggagtt gaagccccggg tcgcgggttt 780  
 atttctccgt aggatcgaca ctttatccag aacacagagc cagcttgcag agctggcatc 840  
 ttccatcggt gttccacttg aatggaaagc agtcgtacac ttgccgtgt gatacagtgg 900  
 tgagttgcga aggctacgta gtgaagaaaa tcaccatcag tcccgggatc acggagaaaa 960  
 ccgtgggata cgcggttaca cacaatagcg agggcttctt gctatgcaaa gttactgaca 1020  
 cagtaaaagg agaacgggta tcgttccctg tgtgcacgta catccggcc accatatgcg 1080  
 atcagatgac tggtataatg gccacggata tatcacctga cgatgcacaa aaacttctgg 1140  
 ttgggctcaa ccagcgaatt gtcattaacg gtaggactaa caggaacacc aacaccatgc 1200  
 aaaattacct tctgccgatc atagcacaag gttcagcaa atgggctaag gagcgcaagg 1260  
 atgatcttga taacgagaaa atgctggta ctagagaacg caagcttacg tatggctgct 1320  
 tgtggcgtt tcgcactaag aaagtacatt cgtttatcg cccacctgga acgcagacct 1380  
 gcgtaaaagt cccagcctct tttagcgctt ttcccatgtc gtccgtatgg acgacctctt 1440  
 tgcccatgtc gctgaggcag aaattgaaac tggcattgca accaaagaag gagaaaaaac 1500  
 tgctgcaggt ctcggaggaa ttagtcatgg aggccaaggc tgctttgag gatgctcagg 1560  
 aggaagccag agcggagaag ctccgagaag cacttccacc attagtggca gacaaaggca 1620  
 tcgaggcagc cgacagaagtt gtctgcgaag tggagggct ccaggcggac atcggagcag 1680

cattagttga aaccccgcgc ggtcacgtaa ggataatacc tcaagcaaat gaccgtatga	1740
tcggacagta tatcgttgtc tcgccaaact ctgtgctgaa gaatgccaaa ctcgcaccag	1800
cgcacccgct agcagatcag gttaagatca taacacactc cggaaagatca ggaaggtacg	1860
cggtcgaacc atacgacgct aaagtactga tgccagcagg aggtgccgta ccatggccag	1920
aattccttagc actgagtgag agcgccacgt tagtgtacaa cgaaagagag tttgtgaacc	1980
gcaaactata ccacattgcc atgcattggcc ccgccaagaa tacagaagag gagcagtaca	2040
aggttacaaa ggcagagctt gcagaaacag agtacgtgtt tgacgtggac aagaagcggtt	2100
gcgttaagaa ggaagaagcc tcaggtctgg tcctctcggg agaactgacc aaccctccct	2160
atcatgagct agctctggag ggactgaaga cccgacactgc ggtcccgat aaggtcgaaa	2220
caataggagt gataggcaca cgggggtcgg gcaagtcagc tattatcaag tcaactgtca	2280
cggcacgaga tcttgttacc agcggaaaga aagaaaattg tcgcgaaatt gaggccgacg	2340
tgctaagact gaggggtatg cagattacgt cgaagacagt agattcggtt atgctcaacg	2400
gatgccacaa agccgtagaa gtgctgtacg ttgacgaagc gttcgcgtgc cacgcaggag	2460
cactacttgc cttgattgtc atcgtcagggc cccgcaagaa ggttagtacta tgcggagacc	2520
ccatgcaatg cggattcttc aacatgatgc aactaaaggt acatttcaat caccctgaaa	2580
aagacatatg caccaagaca ttctacaagt atatctcccg gcgttgcaca cagccagttt	2640
cagctattgt atcgacactg cattacgatg gaaagatgaa aaccacgaac ccgtgcaaga	2700
agaacattga aatcgatatt acaggggcca caaagccgaa gccaggggat atcatcctga	2760
catgtttccg cgggtgggtt aagcaattgc aaatcgacta tcccggacat gaagtaatga	2820
cagccgcggc ctcacaaggg ctaaccagaa aaggagtgtt tgccgtccgg caaaaagtca	2880
atgaaaaccc actgtacgacg atcacatcag agcatgtgaa cgtgttgctc acccgactg	2940
aggacaggtt agtgtggaaa accttgcagg gcgaccatg gattaagcag cccactaaca	3000
tacctaaagg aaactttcag gctactatag aggactggaa agctgaacac aagggaaataa	3060
ttgctgcaat aaacagcccc actccccgtg ccaatccgtt cagctgcaag accaacgttt	3120
gctggcgaa agcattggaa ccgataactag ccacggccgg tatcgactt accgggttgc	3180
agtggagcga actgttccca cagttgcgg atgacaaacc acattcggcc atttacgcct	3240
tagacgtaat ttgcattaag ttttcggca tggacttgac aagcggactg ttttctaaac	3300
agagcatccc actaacgtac catccgcgg attcagcgag gccggtagct cattgggaca	3360

acagcccagg aacccgcaag tatgggtacg atcacgccat tgccgccgaa ctctcccgta 3420  
 gatttccggt gttccagcta gctgggaagg gcacacaact tgattgcag acggggagaa 3480  
 ccagagttat ctctgcacag cataacctgg tcccggtgaa ccgcaatctt cctcacgcct 3540  
 tagtccccga gtacaaggag aagcaacccg gcccggtcaa aaaattcttg aaccagttca 3600  
 aacaccactc agtacttgtg gtatcagagg aaaaaattga agctccccgt aagagaatcg 3660  
 aatggatcgc cccgattggc atagccggtg cagataagaa ctacaacctg gcttcgggt 3720  
 ttccgcccga ggcacggtac gacctggtgt tcatcaacat tggactaaa tacagaaacc 3780  
 accactttca gcagtgcgaa gaccatgcgg cgaccttaaa aacccttcg cgttcggccc 3840  
 tgaattgcct taacccagga ggcaccctcg tggtaagtc ctatggctac gccgaccgca 3900  
 acagtgagga cgtagtcacc gctttgcca gaaagttgt caggggtct gcagcgagac 3960  
 cagattgtgt ctcaagcaat acagaaatgt acctgatttt ccgacaacta gacaacagcc 4020  
 gtacacggca attcaccccg caccatctga attgcgtgat ttctccgtg tatgagggtt 4080  
 caagagatgg agttggagcc ggcgcgtcat accgcaccaa aaggagaat attgctgact 4140  
 gtcaagagga agcagttgtc aacgcagcca atccgctggg tagaccaggc gaaggagtct 4200  
 gccgtgccat ctataaacgt tggccgacca gtttaccga ttccagccacg gagacaggca 4260  
 ccgcaagaat gactgtgtgc ctaggaaaga aagtgttcca cgcggcggc cctgatttcc 4320  
 ggaagcaccc agaagcagaa gccttggaaat tgctacaaaa cgcctaccat gcagtggcag 4380  
 acttagtaaa tgaacataac atcaagtctg tcgccattcc actgctatct acaggcattt 4440  
 acgcagccgg aaaagaccgc cttgaagtat cacttaactg cttgacaacc ggcgttagaca 4500  
 gaactgacgc ggacgtaacc atctattgcc tggataagaa gtggaaaggaa agaatcgacg 4560  
 cggcactcca acttaaggag tctgttaacag agctgaagga tgaagatatg gagatcgacg 4620  
 atgagttatgt atggatccat ccagacagtt gcttgaaggg aagaaaggaa ttctgttacta 4680  
 caaaaaggaaa attgtattcg tacttcgaag gcaccaaatt ccatcaagca gcaaaagaca 4740  
 tggcggagat aaaggctctg ttccctaattg accaggaaag taatgaacaa ctgtgtgcct 4800  
 acatattggg tgagaccatg gaagcaatcc gcgaaaagtg cccggcgcac cataaccctgt 4860  
 cgtctagccc gccccaaacg ttgcgtgcc tttgcattgtc tgccatgcac ccagaaagg 4920  
 tccacagact tagaagcaat aacgtcaaag aagttacagt atgctccctcc accccccttc 4980  
 ctaagcacaat aatataagaat gttcagaagg ttctgtgcac gaaagttagtc ctgtttaatc 5040  
 cgcacactcc cgcattcggtt cccggccgtt agtacataga agtgccagaa cagcctaccg 5100

ctccctcctgc acaggccgag gaggccccg aagttgtac gacaccgtca ccatctacag 5160  
 ctgataacac ctcgcttgat gtcacagaca tctcactgga tatggatgac agtagcgaag 5220  
 gctcacttt ttcgagctt agcggatcg acaactctat tactagtatg gacagttgg 5280  
 cgtcaggacc tagttcacta gagatagtag accgaaggca ggtggtggg gctgacgttc 5340  
 atgccgtcca agagcctgcc cctattccac cgccaaggct aaagaagatg gcccgccctgg 5400  
 cagcggcaag aaaagagccc actccaccgg caagcaatag ctctgagtcc ctccacctct 5460  
 ctttggtgg ggtatccatg tccctcgat caatttcga cggagagacg gcccggcagg 5520  
 cagcggtaca acccctggca acaggcccc a cgatgtgcc tatgtcttc ggatcgttt 5580  
 ccgacggaga gattgatgag ctgagccgca gagtaactga gtccgaaccc gtcctgttt 5640  
 gatcatttga accggggcaa gtgaactcaa ttatatcgcc cccatcgcc gtatctttc 5700  
 cactacgcaa gcagagacgt agacgcagga gcaggaggac tgaatactga ctaaccgggg 5760  
 taggtgggta catatttcg acggacacag gccctggca cttgcaaaag aagtccgttc 5820  
 tgcagaacca gcttacagaa ccgacctgg agcgcaatgt cctggaaaga attcatgccc 5880  
 cggtgctcga cacgtcgaaa gaggaacaac tcaaactcag gtaccagatg atgcccaccc 5940  
 aagccaacaa aagtaggtac cagtctcgta aagtagaaaa tcagaaagcc ataaccactg 6000  
 agcgactact gtcaggacta cgactgtata actctgccac agatcagccca gaatgctata 6060  
 agatcaccta tccgaaacca ttgtactcca gtagcgtacc ggcgaactac tccgatccac 6120  
 agttcgctgt agctgtctgt aacaactatc tgcgtgagaa ctatccgaca gtagcatctt 6180  
 atcagattac tgacgagttac gatgcttact tggatatggt agacgggaca gtcgcctgccc 6240  
 tggatactgc aaccttctgc cccgctaagc ttagaagtta cccgaaaaaaaa catgagtata 6300  
 gagccccgaa tatccgcagt gcggttccat cagcgatgca gaacacgcta caaaatgtgc 6360  
 tcattgccgc aactaaaaga aattgcaacg tcacgcagat gcgtgaactg ccaacactgg 6420  
 actcagcgac attcaatgtc gaatgcttgc gaaaatatgc atgtaatgac gagtattggg 6480  
 aggagttcgc tcggaaggca attaggatta ccactgagtt tgtcaccgca tatgtagcta 6540  
 gactgaaagg ccctaaggcc gccgcactat ttgcaaagac gtataatttg gtcccatcg 6600  
 aagaagtgcc tatggataga ttcgtcatgg acatgaaaag agacgtgaaa gttacaccag 6660  
 gcacgaaaca cacagaagaa agaccgaaag tacaagtgtat acaagccgca gaaccctgg 6720  
 cgactgctta ctatgcggg attcaccggg aatttagtgcg taggcttacg gccgtcttgc 6780

ttccaaacat tcacacgctt tttgacatgt cggcggagga ttttcatgca atcatagcag	6840
aacacttcaa gcaaggcgac ccggtaactgg agacggatat cgcatttc gacaaaagcc	6900
aagacgacgc tatggcgta accggctctga tgatcttggg ggacctgggt gtggatcaac	6960
cactactcga cttgatcgag tgcccttg gagaaatatc atccacccat ctacctacgg	7020
gtactcgtt taaattcggg gcgtatgtga aatccggat gttccctcaca cttttgtca	7080
acacagttt gaatgtcggt atcgccagca gagtactaga agagcggctt aaaacgtcca	7140
gatgtgcagc gttcattggc gacgacaaca tcatacatgg agtagtatct gacaaagaaa	7200
tggctgagag gtgcgccacc tggctcaaca tggaggttaa gatcatcgac gcagtcatcg	7260
gtgagagacc accttacttc tgcggcgat ttatcttgcg agattcggtt acttccacag	7320
cgtgccgcgt ggccggatccc ctgaaaaggc tggtaaagtt gggtaaaccg ctccagccg	7380
acgacgagca agacgaagac agaagacgacg ctctgtatcg tggaaacaaag gcgtggttt	7440
gagtaggtat aacaggcact ttagcagtgg ccgtgacgac ccgtatgag gttagacaata	7500
ttacacctgt cctactggca ttgagaactt ttgcccaagag caaaagagca ttccaagcca	7560
tcagagggga aataaagcat ctctacggtg gtcctaaata gtcagcatag tacatttcat	7620
ctgactaata ctacaacacc accaccatga atagaggatt cttaacatg ctcggccgccc	7680
gcccttcccc ggcccccaact gccatgtgga ggccgcggag aaggaggcag gcggcccccga	7740
tgctgcccgg caacgggctg gcttctcaaa tccagcaact gaccacagcc gtcagtgcgg	7800
tagtcattgg acaggcaact agacctaacc cccacgtcc acgcccggca ccgcgcggaga	7860
agaagcagggc gcccaagcaa ccaccgaagc cgaagaaacc aaaaacgcag gagaagaaga	7920
agaagcaacc tgcaaaaccc aaacccggaa agagacagcg catggcactt aagttggagg	7980
ccgacagatt gttcgacgatc aagaacgagg acggagatgt catcgccac gcactggcca	8040
tggaaaggaaa ggtatgaaa cctctgcacg tggaaaggaaac catcgaccac cctgtgctat	8100
caaagctcaa atttaccaag tcgtcagcat acgacatggg gttcgacag ttgccagtca	8160
acatgagaag tgaggcattc acctacacca gtgaacacccc cgaaggattc tataactggc	8220
accacggagc ggtgcagttt agtggaggtt gatttaccat ccctcgccggaa gtggaggcag	8280
gaggagacag cggtcgtccg atcatggata actccggatcg ggttgcgcgcg atagtcctcg	8340
gtggcgctga tggaaacca cgaactgccc ttccggcgat cacctggaaat agtaaaggaa	8400
agacaattaa gacgaccccg gaagggacag aagagtgg tcc gca gca cca ctg gtc	8456
	Ser Ala Ala Pro Leu Val
1	5

acg gca atg tgt ttg ctc gga aat gtg agc ttc cca tgc gac cgc ccg	10	15	20	8504
Thr Ala Met Cys Leu Leu Gly Asn Val Ser Phe Pro Cys Asp Arg Pro				
ccc aca tgc tat acc cgc gaa cct tcc aga gcc ctc gac atc ctt gaa	25	30	35	8552
Pro Thr Cys Tyr Thr Arg Glu Pro Ser Arg Ala Leu Asp Ile Leu Glu				
gag aac gtg aac cat gag gcc tac gat acc ctg ctc aat gcc ata ttg	40	45	50	8600
Glu Asn Val Asn His Glu Ala Tyr Asp Thr Leu Leu Asn Ala Ile Leu				
cgg tgc gga tcg tct ggc aga agc aaa aga agc gtc att gac gac ttt	55	60	65	8648
Arg Cys Gly Ser Ser Gly Arg Ser Lys Arg Ser Val Ile Asp Asp Phe				
acc ctg acc agc ccc tac ttg ggc aca tgc tcg tac tgc cac cat act	75	80	85	8696
Thr Leu Thr Ser Pro Tyr Leu Gly Thr Cys Ser Tyr Cys His His Thr				
gta ccg tgc ttc agc cct gtt aag atc gag cag gtc tgg gac gaa gcg	90	95	100	8744
Val Pro Cys Phe Ser Pro Val Lys Ile Glu Gln Val Trp Asp Glu Ala				
gac gat aac acc ata cgc ata cag act tcc gcc cag ttt gga tac gac	105	110	115	8792
Asp Asp Asn Thr Ile Arg Ile Gln Thr Ser Ala Gln Phe Gly Tyr Asp				
caa agc gga gca gca agc gca aac aag tac cgc tac atg tcg ctt aag	120	125	130	8840
Gln Ser Gly Ala Ala Ser Ala Asn Lys Tyr Arg Tyr Met Ser Leu Lys				
cag gat cac acc gtt aaa gaa ggc acc atg gat gac atc aag att agc	135	140	145	8888
Gln Asp His Thr Val Lys Glu Gly Thr Met Asp Asp Ile Lys Ile Ser				
acc tca gga ccg tgt aga agg ctt agc tac aaa gga tac ttt ctc ctc	155	160	165	8936
Thr Ser Gly Pro Cys Arg Arg Leu Ser Tyr Lys Gly Tyr Phe Leu Leu				
gca aaa tgc cct cca ggg gac agc gta acg gtt agc ata gtg agt agc	170	175	180	8984
Ala Lys Cys Pro Pro Gly Asp Ser Val Thr Val Ser Ile Val Ser Ser				
aac tca gca acg tca tgt aca ctg gcc cgc aag ata aaa cca aaa ttc	185	190	195	9032
Asn Ser Ala Thr Ser Cys Thr Leu Ala Arg Lys Ile Lys Pro Lys Phe				
gtg gga cgg gaa aaa tat gat cta cct ccc gtt cac ggt aaa aaa att	200	205	210	9080
Val Gly Arg Glu Lys Tyr Asp Leu Pro Pro Val His Gly Lys Lys Ile				
cct tgc aca gtg tac gac cgt ctg aaa gaa aca act gca ggc tac atc	215	220	225	9128
Pro Cys Thr Val Tyr Asp Arg Leu Lys Glu Thr Thr Ala Gly Tyr Ile				

act atg cac agg ccg aga ccg cac gct tat aca tcc tac ctg gaa gaa	9176
Thr Met His Arg Pro Arg Pro His Ala Tyr Thr Ser Tyr Leu Glu Glu	
235 240 245	
tca tca ggg aaa gtt tac gca aag ccg cca tct ggg aag aac att acg	9224
Ser Ser Gly Lys Val Tyr Ala Lys Pro Pro Ser Gly Lys Asn Ile Thr	
250 255 260	
tat gag tgc aag tgc ggc gac tac aag acc gga acc gtt tcg acc cgc	9272
Tyr Glu Cys Lys Cys Gly Asp Tyr Lys Thr Gly Thr Val Ser Thr Arg	
265 270 275	
acc gaa atc act ggt tgc acc gcc atc aag cag tgc gtc gcc tat aag	9320
Thr Glu Ile Thr Gly Cys Thr Ala Ile Lys Gln Cys Val Ala Tyr Lys	
280 285 290	
agc gac caa acg aag tgg gtc ttc aac tca ccg gac ttg atc aga cat	9368
Ser Asp Gln Thr Lys Trp Val Phe Asn Ser Pro Asp Leu Ile Arg His	
295 300 305 310	
gac gac cac acg gcc caa ggg aaa ttg cat ttg cct ttc aag ttg atc	9416
Asp Asp His Thr Ala Gln Gly Lys Leu His Leu Pro Phe Lys Leu Ile	
315 320 325	
ccg agt acc tgc atg gtc cct gtt gcc cac gcg ccg aat gta ata cat	9464
Pro Ser Thr Cys Met Val Pro Val Ala His Ala Pro Asn Val Ile His	
330 335 340	
ggc ttt aaa cac atc agc ctc caa tta gat aca gac cac ttg aca ttg	9512
Gly Phe Lys His Ile Ser Leu Gln Leu Asp Thr Asp His Leu Thr Leu	
345 350 355	
ctc acc acc agg aga cta ggg gca aac ccg gaa cca acc act gaa tgg	9560
Leu Thr Thr Arg Arg Leu Gly Ala Asn Pro Glu Pro Thr Thr Glu Trp	
360 365 370	
atc gtc gga aag acg gtc aga aac ttc acc gtc gac cga gat ggc ctg	9608
Ile Val Gly Lys Thr Val Arg Asn Phe Thr Val Asp Arg Asp Gly Leu	
375 380 385 390	
gaa tac ata tgg gga aat cat gag cca gtg agg gtc tat gcc caa gag	9656
Glu Tyr Ile Trp Gly Asn His Glu Pro Val Arg Val Tyr Ala Gln Glu	
395 400 405	
tca gca cca gga gac cct cac gga tgg cca cac gaa ata gta cag cat	9704
Ser Ala Pro Gly Asp Pro His Gly Trp Pro His Glu Ile Val Gln His	
410 415 420	
tac tac cat cgc cat cct gtg tac acc atc tta gcc gtc gca tca gct	9752
Tyr Tyr His Arg His Pro Val Tyr Thr Ile Leu Ala Val Ala Ser Ala	
425 430 435	
acc gtg gcg atg atg att ggc gta act gtt gca gtg tta tgt gcc tgt	9800
Thr Val Ala Met Met Ile Gly Val Thr Val Ala Val Leu Cys Ala Cys	
440 445 450	
aaa gcg cgc cgt gag tgc ctg acg cca tac gcc ctg gcc cca aac gcc	9848

Lys	Ala	Arg	Arg	Glu	Cys	Leu	Thr	Pro	Tyr	Ala	Leu	Ala	Pro	Asn	Ala	
455				460				465		470						
gta	atc	cca	act	tcg	ctg	gca	ctc	ttg	tgc	tgc	gtt	agg	tcg	gcc	aat	9896
Val	Ile	Pro	Thr	Ser	Leu	Ala	Leu	Leu	Cys	Cys	Val	Arg	Ser	Ala	Asn	
				475				480		485						
gct	gaaacgttca	ccgagaccat	gagttacttg	tggtcgaaca	gtcagccgtt											9949
Ala																
cttctgggtc	cagttgtgca	tacctttggc	cgctttcatc	gttctaatgc	gctgctgctc											10009
ctgctgcctg	ccttttttag	tggttgccgg	cgcttacctg	gcgaaggtag	acgcctacga											10069
acatgcgacc	actgttccaa	atgtgccaca	gataccgtat	aaggcacttg	ttgaaagggc											10129
agggtatgcc	ccgctcaatt	tggagatcac	tgtcatgtcc	tcggaggttt	tgccttccac											10189
caaccaagag	tacattacct	gcaaattcac	cactgtggtc	ccctccccaa	aaatcaaatg											10249
ctgcggctcc	ttggaatgtc	agccggccgc	tcatgcagac	tatacctgca	aggtcttcgg											10309
aggggtctac	ccctttatgt	ggggaggagc	gcaatgtttt	tgcgacagtg	agaacagcca											10369
gatgagttag	gcgtacgtcg	aattgtcagc	agattgcgcg	tctgaccacg	cgcaggcgat											10429
taaggtgcac	actgccgcga	tgaaagttag	actgcgtatt	gtgtacggga	acactaccag											10489
tttccttagat	gtgtacgtga	acggagtcac	accaggaacg	tctaaagact	tgaaagtcat											10549
agctggacca	atttcagcat	cgtttacgcc	attcgatcat	aaggtcgta	tccatcgccgg											10609
cctgggtgtac	aactatgact	tcccggata	tggagcgtat	aaaccaggag	cgtttggaga											10669
cattcaagct	accccttga	ctagcaagga	tctcatgcc	agcacagaca	ttaggctact											10729
caagccttcc	gccaagaacg	tgcgttccc	gtacacgcag	gcctcatcag	gatttggat											10789
gtggaaaaac	aactcaggcc	gcccaactgca	ggaaaccgca	ccttcgggt	gtaagattgc											10849
agtaaatccg	ctccgagcgg	tggactgttc	atacggaaac	attccattt	ctattgacat											10909
cccgaaacgt	gcctttatca	ggacatcaga	tgcaccactg	gtctcaacag	tcaaattgtga											10969
agtcagttag	tgcacttatt	cagcagactt	cggcgggatg	gccaccctgc	agtatgtatc											11029
cgaccgcgaa	ggtcaatgcc	ccgtacattc	gcattcgagc	acagcaactc	tccaagagtc											11089
gacagtacat	gtcctggaga	aaggagcggt	gacagtacac	tttagcaccg	cgagtccaca											11149
ggcgaacttt	atcgttatgc	tgtgtggaa	gaagacaaca	tgcaatgcag	aatgtaaacc											11209
accagctgac	catatcgtga	gcaccccgca	caaaaatgac	caagaatttc	aagccgccat											11269
ctcaaaaaca	tcatggagtt	ggctgtttgc	cctttcggc	ggcgctcgt	cgctattaat											11329

tataggactt atgattttg cttgcagcat gatgctgact agcacacgaa gatgaccgct 11389  
 acgccccaat gatccgacca gcaaaactcg atgtacttcc gaggaactga tgtgcataat 11449  
 gcatcaggct ggtacattag atccccgctt accgcgggca atatagcaac actaaaaact 11509  
 cgatgtactt ccgaggaagc gcagtgcata atgctgca 11569  
 tattaaccat ttatctagcg gacgccaaaa actcaatgta tttctgagga agcgtggtgc 11629  
 ataatgccac gcagcgtctg cataactttt attatttctt ttattaatca acaaaatttt 11689  
 gtttttaaca ttcc 11703

<210> 16  
 <211> 64  
 <212> PRT  
 <213> Sindbis virus

<400> 16

Ser Ala Ala Pro Leu Val Thr Ala Met Cys Leu Leu Gly Asn Val Ser  
 1 5 10 15

Phe Pro Cys Asp Arg Pro Pro Thr Cys Tyr Thr Arg Glu Pro Ser Arg  
 20 25 30

Ala Leu Asp Ile Leu Glu Glu Asn Val Asn His Glu Ala Tyr Asp Thr  
 35 40 45

Leu Leu Asn Ala Ile Leu Arg Cys Gly Ser Ser Gly Arg Ser Lys Arg  
 50 55 60

<210> 17  
 <211> 423  
 <212> PRT  
 <213> Sindbis virus

<400> 17

Ser Val Ile Asp Asp Phe Thr Leu Thr Ser Pro Tyr Leu Gly Thr Cys  
 1 5 10 15

Ser Tyr Cys His His Thr Val Pro Cys Phe Ser Pro Val Lys Ile Glu  
 20 25 30

Gln Val Trp Asp Glu Ala Asp Asp Asn Thr Ile Arg Ile Gln Thr Ser  
 35 40 45

Ala Gln Phe Gly Tyr Asp Gln Ser Gly Ala Ala Ser Ala Asn Lys Tyr  
50 55 60

Arg Tyr Met Ser Leu Lys Gln Asp His Thr Val Lys Glu Gly Thr Met  
65 70 75 80

Asp Asp Ile Lys Ile Ser Thr Ser Gly Pro Cys Arg Arg Leu Ser Tyr  
85 90 95

Lys Gly Tyr Phe Leu Leu Ala Lys Cys Pro Pro Gly Asp Ser Val Thr  
100 105 110

Val Ser Ile Val Ser Ser Asn Ser Ala Thr Ser Cys Thr Leu Ala Arg  
115 120 125

Lys Ile Lys Pro Lys Phe Val Gly Arg Glu Lys Tyr Asp Leu Pro Pro  
130 135 140

Val His Gly Lys Lys Ile Pro Cys Thr Val Tyr Asp Arg Leu Lys Glu  
145 150 155 160

Thr Thr Ala Gly Tyr Ile Thr Met His Arg Pro Arg Pro His Ala Tyr  
165 170 175

Thr Ser Tyr Leu Glu Glu Ser Ser Gly Lys Val Tyr Ala Lys Pro Pro  
180 185 190

Ser Gly Lys Asn Ile Thr Tyr Glu Cys Lys Cys Gly Asp Tyr Lys Thr  
195 200 205

Gly Thr Val Ser Thr Arg Thr Glu Ile Thr Gly Cys Thr Ala Ile Lys  
210 215 220

Gln Cys Val Ala Tyr Lys Ser Asp Gln Thr Lys Trp Val Phe Asn Ser  
225 230 235 240

Pro Asp Leu Ile Arg His Asp Asp His Thr Ala Gln Gly Lys Leu His  
245 250 255

Leu Pro Phe Lys Leu Ile Pro Ser Thr Cys Met Val Pro Val Ala His  
260 265 270

Ala Pro Asn Val Ile His Gly Phe Lys His Ile Ser Leu Gln Leu Asp

275	280	285
Thr Asp His Leu Thr Leu Leu Thr Thr Arg Arg Leu Gly Ala Asn Pro		
290	295	300
Glu Pro Thr Thr Glu Trp Ile Val Gly Lys Thr Val Arg Asn Phe Thr		
305	310	315
320		
Val Asp Arg Asp Gly Leu Glu Tyr Ile Trp Gly Asn His Glu Pro Val		
325	330	335
Arg Val Tyr Ala Gln Glu Ser Ala Pro Gly Asp Pro His Gly Trp Pro		
340	345	350
His Glu Ile Val Gln His Tyr Tyr His Arg His Pro Val Tyr Thr Ile		
355	360	365
Leu Ala Val Ala Ser Ala Thr Val Ala Met Met Ile Gly Val Thr Val		
370	375	380
Ala Val Leu Cys Ala Cys Lys Ala Arg Arg Glu Cys Leu Thr Pro Tyr		
385	390	395
400		
Ala Leu Ala Pro Asn Ala Val Ile Pro Thr Ser Leu Ala Leu Leu Cys		
405	410	415
Cys Val Arg Ser Ala Asn Ala		
420		

<210> 18  
<211> 51  
<212> DNA  
<213> Artificial sequence

<220>  
<223> Linker

<220>  
<221> CDS  
<222> (1)..(51)

<400> 18  
aga tct ggt ggc ggt ggc tcg ggc ggt ggt ggg tcg ggt ggc ggc gga 48  
Arg Ser Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly Gly  
1 5 10 15

tct  
Ser

51

<210> 19  
<211> 17  
<212> PRT  
<213> Artificial sequence

<220>  
<223> Linker

<400> 19

Arg Ser Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly Gly  
1 5 10 15

Ser

<210> 20  
<211> 45  
<212> DNA  
<213> Artificial sequence

<220>  
<223> Linker

<220>  
<221> CDS  
<222> (1)..(45)

<400> 20  
ggg ggc ggt ggc tcg ggc ggt ggt ggg tcg ggt ggc ggc gga tct  
Gly Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly Gly Ser  
1 5 10 15

45

<210> 21  
<211> 15  
<212> PRT  
<213> Artificial sequence

<220>  
<223> Linker

<400> 21

Gly Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly Gly Ser  
1 5 10 15

<210> 22

<211> 1050  
 <212> DNA  
 <213> Bluetongue virus 10

&lt;220&gt;

<221> CDS  
 <222> (1)..(1050)  
 <223> VP7 gene

&lt;400&gt; 22

atg gac act atc gcc gca aga gca ctc act gtg atg cga gca tgt gct 48  
 Met Asp Thr Ile Ala Ala Arg Ala Leu Thr Val Met Arg Ala Cys Ala  
 1 5 10 15

acg ctt caa gag gca aga att gtg ttg gaa gcc aat gtg atg gaa att 96  
 Thr Leu Gln Glu Ala Arg Ile Val Leu Glu Ala Asn Val Met Glu Ile  
 20 25 30

ttg ggg ata gct atc aat agg tac aat gga ctc act tta cga gga gtg 144  
 Leu Gly Ile Ala Ile Asn Arg Tyr Asn Gly Leu Thr Leu Arg Gly Val  
 35 40 45

acg atg cgc ccg acc tcg tta gca caa aga aat gag atg ttt ttt atg 192  
 Thr Met Arg Pro Thr Ser Leu Ala Gln Arg Asn Glu Met Phe Phe Met  
 50 55 60

tgt ttg gat atg atg ctg tct gct gct ggg ata aat gtt gga ccg ata 240  
 Cys Leu Asp Met Met Leu Ser Ala Ala Gly Ile Asn Val Gly Pro Ile  
 65 70 75 80

tcg cca gac tat act caa cat atg gct acg att ggt gta cta gca aca 288  
 Ser Pro Asp Tyr Thr Gln His Met Ala Thr Ile Gly Val Leu Ala Thr  
 85 90 95

ccg gaa ata cct ttt aca acg gaa gcg gcg aat gaa ata gca cga gtg 336  
 Pro Glu Ile Pro Phe Thr Thr Glu Ala Ala Asn Glu Ile Ala Arg Val  
 100 105 110

act ggg gag act tcg aca tgg ggg cca gcg cgt cag cct tat ggt ttc 384  
 Thr Gly Glu Thr Ser Thr Trp Gly Pro Ala Arg Gln Pro Tyr Gly Phe  
 115 120 125

ttc ctt gaa act gag gaa acc ttc caa cca ggg agg tgg ttc atg cgc 432  
 Phe Leu Glu Thr Glu Glu Thr Phe Gln Pro Gly Arg Trp Phe Met Arg  
 130 135 140

gcc gct caa gca gta act gca gta gtg tgc ggt ccg gat atg att caa 480  
 Ala Ala Gln Ala Val Thr Ala Val Val Cys Gly Pro Asp Met Ile Gln  
 145 150 155 160

gtg tca ctt aat gct gga gcg aga gga gat gta caa cag ata ttt cag 528  
 Val Ser Leu Asn Ala Gly Ala Arg Gly Asp Val Gln Gln Ile Phe Gln  
 165 170 175

ggt cgt aat gat ccc atg atg ata tat tta gtg tgg agg aga atc gaa 576  
 Gly Arg Asn Asp Pro Met Met Ile Tyr Leu Val Trp Arg Arg Ile Glu

180	185	190	
aac ttt gcg atg gcg caa ggt aat tca cag caa act caa gcg ggt gtg Asn Phe Ala Met Ala Gln Gly Asn Ser Gln Gln Thr Gln Ala Gly Val 195	200	205	624
act gtc agt gtt ggt gga gtt gac atg agg gcg gga cgc att ata gcg Thr Val Ser Val Gly Gly Val Asp Met Arg Ala Gly Arg Ile Ile Ala 210	215	220	672
tgg gat gga cag gcc gcg ctg cat gtg cat aat ccg aca caa cag aat Trp Asp Gly Gln Ala Ala Leu His Val His Asn Pro Thr Gln Gln Asn 225	230	235	720
gcg atg gtg caa ata cag gtt gtg ttc tat ata tct atg gat aaa act Ala Met Val Gln Ile Gln Val Val Phe Tyr Ile Ser Met Asp Lys Thr 245	250	255	768
tta aac cag tac ccc gct ttg act gct gag att ttc aat gtt tac agc Leu Asn Gln Tyr Pro Ala Leu Thr Ala Glu Ile Phe Asn Val Tyr Ser 260	265	270	816
ttc agg gac cac aca tgg cat ggg cta aga acg gcg ata tta aac aga Phe Arg Asp His Thr Trp His Gly Leu Arg Thr Ala Ile Leu Asn Arg 275	280	285	864
acc aca ctg cca aac atg ctg cca cca atc ttc cca cca aat gat cga Thr Thr Leu Pro Asn Met Leu Pro Pro Ile Phe Pro Pro Asn Asp Arg 290	295	300	912
gat agc atc tta act ctt cta ctt tta tct aca ctt gct gat gtt tac Asp Ser Ile Leu Thr Leu Leu Leu Ser Thr Leu Ala Asp Val Tyr 305	310	315	960
act gtt tta agg cca gag ttt gcg att cac ggc gta aat ccg atg cca Thr Val Leu Arg Pro Glu Phe Ala Ile His Gly Val Asn Pro Met Pro 325	330	335	1008
ggg ccg ctc aca cgt gct att gcg cgc gcc gcc tat gtg tag Gly Pro Leu Thr Arg Ala Ile Ala Arg Ala Ala Tyr Val 340	345		1050

<210> 23  
 <211> 349  
 <212> PRT  
 <213> Bluetongue virus 10  
  
 <400> 23

Met Asp Thr Ile Ala Ala Arg Ala Leu Thr Val Met Arg Ala Cys Ala  
 1 5 10 15

Thr Leu Gln Glu Ala Arg Ile Val Leu Glu Ala Asn Val Met Glu Ile  
 20 25 30

Leu Gly Ile Ala Ile Asn Arg Tyr Asn Gly Leu Thr Leu Arg Gly Val  
35 40 45

Thr Met Arg Pro Thr Ser Leu Ala Gln Arg Asn Glu Met Phe Phe Met  
50 55 60

Cys Leu Asp Met Met Leu Ser Ala Ala Gly Ile Asn Val Gly Pro Ile  
65 70 75 80

Ser Pro Asp Tyr Thr Gln His Met Ala Thr Ile Gly Val Leu Ala Thr  
85 90 95

Pro Glu Ile Pro Phe Thr Thr Glu Ala Ala Asn Glu Ile Ala Arg Val  
100 105 110

Thr Gly Glu Thr Ser Thr Trp Gly Pro Ala Arg Gln Pro Tyr Gly Phe  
115 120 125

Phe Leu Glu Thr Glu Glu Thr Phe Gln Pro Gly Arg Trp Phe Met Arg  
130 135 140

Ala Ala Gln Ala Val Thr Ala Val Val Cys Gly Pro Asp Met Ile Gln  
145 150 155 160

Val Ser Leu Asn Ala Gly Ala Arg Gly Asp Val Gln Gln Ile Phe Gln  
165 170 175

Gly Arg Asn Asp Pro Met Met Ile Tyr Leu Val Trp Arg Arg Ile Glu  
180 185 190

Asn Phe Ala Met Ala Gln Gly Asn Ser Gln Gln Thr Gln Ala Gly Val  
195 200 205

Thr Val Ser Val Gly Gly Val Asp Met Arg Ala Gly Arg Ile Ile Ala  
210 215 220

Trp Asp Gly Gln Ala Ala Leu His Val His Asn Pro Thr Gln Gln Asn  
225 230 235 240

Ala Met Val Gln Ile Gln Val Val Phe Tyr Ile Ser Met Asp Lys Thr  
245 250 255

Leu Asn Gln Tyr Pro Ala Leu Thr Ala Glu Ile Phe Asn Val Tyr Ser  
260 265 270

Phe Arg Asp His Thr Trp His Gly Leu Arg Thr Ala Ile Leu Asn Arg  
275 280 285

Thr Thr Leu Pro Asn Met Leu Pro Pro Ile Phe Pro Pro Asn Asp Arg  
290 295 300

Asp Ser Ile Leu Thr Leu Leu Leu Leu Ser Thr Leu Ala Asp Val Tyr  
305 310 315 320

Thr Val Leu Arg Pro Glu Phe Ala Ile His Gly Val Asn Pro Met Pro  
325 330 335

Gly Pro Leu Thr Arg Ala Ile Ala Arg Ala Ala Tyr Val  
340 345

<210> 24  
<211> 585  
<212> DNA  
<213> *Coccidioides immitis*

<220>  
<221> CDS  
<222> (1) .. (585)  
<223> Ag2/PRA gene

<400> 24  
atg cag ttc tct cac gct ctc atc gct ctc gtc gct gcc ggc ctc gcc 48  
Met Gln Phe Ser His Ala Leu Ile Ala Leu Val Ala Ala Gly Leu Ala  
1 5 10 15

agt gcc cag ctc cca gac atc cca cct tgc gct ctc aac tgc ttc gtt 96  
 Ser Ala Gln Leu Pro Asp Ile Pro Pro Cys Ala Leu Asn Cys Phe Val  
 . 20 25 30

cac tgc tcc aag cct gag cta cca gga cag atc act cct tgc gtt gag 192  
 His Cys Ser Lys Pro Glu Leu Pro Gly Gln Ile Thr Pro Cys Val Glu  
 50 55 60

gag gcc tgc cct ctc gac gcc cgt atc tcc gtc tcc aac atc gtc gtt	240		
Glu Ala Cys Pro Leu Asp Ala Arg Ile Ser Val Ser Asn Ile Val Val			
65	70	75	80

gac cag tgc tcc aag gcc ggt gtc cca att gac atc cca cca gtt gac 288

Asp	Gln	Cys	Ser	Lys	Ala	Gly	Val	Pro	Ile	Asp	Ile	Pro	Pro	Val	Asp	
				85				90				95				
acc	acc	gcc	gct	ccc	gag	cca	tcc	gag	acc	gct	gag	ccc	acc	gct	gag	336
Thr	Thr	Ala	Ala	Pro	Glu	Pro	Ser	Glu	Thr	Ala	Glu	Pro	Thr	Ala	Glu	
				100				105				110				
cca	acc	gag	gag	ccc	act	gcc	gag	cct	acc	gct	gag	ccc	acc	gct	gag	384
Pro	Thr	Glu	Glu	Pro	Thr	Ala	Glu	Pro	Thr	Ala	Glu	Pro	Thr	Ala	Glu	
				115				120				125				
ccg	act	cat	gag	ccc	acc	gag	gag	ccc	act	gcc	gtc	cca	acc	ggc	act	432
Pro	Thr	His	Glu	Pro	Thr	Glu	Glu	Pro	Thr	Ala	Val	Pro	Thr	Gly	Thr	
				130				135			140					
ggc	ggt	ggt	gtc	ccc	act	ggc	acc	ggt	tcc	ttc	acc	gtc	act	ggc	aga	480
Gly	Gly	Gly	Val	Pro	Thr	Gly	Thr	Gly	Ser	Phe	Thr	Val	Thr	Gly	Arg	
				145				150			155				160	
cca	act	gcc	tcc	acc	cca	gct	gag	ttc	cca	ggt	gct	ggc	tcc	aac	gtc	528
Pro	Thr	Ala	Ser	Thr	Pro	Ala	Glu	Phe	Pro	Gly	Ala	Gly	Ser	Asn	Val	
				165				170			175					
cgt	gcc	agc	gtt	ggc	ggc	att	gct	gct	ctc	ctc	ggt	ctc	gct	gcc		576
Arg	Ala	Ser	Val	Gly	Gly	Ile	Ala	Ala	Leu	Leu	Gly	Leu	Ala	Ala		
				180				185			190					
tac	ctg	taa														585
Tyr	Leu															

<210> 25  
<211> 194  
<212> PRT  
<213> Coccidioides immitis

<400> 25

Met Gln Phe Ser His Ala Leu Ile Ala Leu Val Ala Ala Gly Leu Ala  
1 5 10 15

Ser Ala Gln Leu Pro Asp Ile Pro Pro Cys Ala Leu Asn Cys Phe Val  
20 25 30

Glu Ala Leu Gly Asn Asp Gly Cys Thr Arg Leu Thr Asp Phe Lys Cys  
35 40 45

His Cys Ser Lys Pro Glu Leu Pro Gly Gln Ile Thr Pro Cys Val Glu  
50 55 60

Glu Ala Cys Pro Leu Asp Ala Arg Ile Ser Val Ser Asn Ile Val Val  
65 70 75 80

Asp Gln Cys Ser Lys Ala Gly Val Pro Ile Asp Ile Pro Pro Val Asp  
 85 90 95

Thr Thr Ala Ala Pro Glu Pro Ser Glu Thr Ala Glu Pro Thr Ala Glu  
 100 105 110

Pro Thr Glu Glu Pro Thr Ala Glu Pro Thr Ala Glu Pro Thr Ala Glu  
 115 120 125

Pro Thr His Glu Pro Thr Glu Glu Pro Thr Ala Val Pro Thr Gly Thr  
 130 135 140

Gly Gly Gly Val Pro Thr Gly Thr Gly Ser Phe Thr Val Thr Gly Arg  
 145 150 155 160

Pro Thr Ala Ser Thr Pro Ala Glu Phe Pro Gly Ala Gly Ser Asn Val  
 165 170 175

Arg Ala Ser Val Gly Gly Ile Ala Ala Ala Leu Leu Gly Leu Ala Ala  
 180 185 190

Tyr Leu

<210> 26  
 <211> 906  
 <212> DNA  
 <213> *Streptococcus pneumoniae*

<220>  
 <221> CDS  
 <222> (1)..(906)  
 <223> *PspA* gene

<400> 26  
 gaa gaa tct ccc gta gcc agt cag tct aaa gct gag aaa gac tat gat 48  
 Glu Glu Ser Pro Val Ala Ser Gln Ser Lys Ala Glu Lys Asp Tyr Asp  
 1 5 10 15

gca gcg aag aaa gat gct aag aat gcg aaa aaa gca gta gaa gat gct 96  
 Ala Ala Lys Lys Asp Ala Lys Asn Ala Lys Lys Ala Val Glu Asp Ala  
 20 25 30

caa aag gct tta gat gat gca aaa gct gct cag aaa aaa tat gac gag 144  
 Gln Lys Ala Leu Asp Asp Ala Lys Ala Ala Gln Lys Lys Tyr Asp Glu  
 35 40 45

gat cag aag aaa act gag gag aaa gcc gcg cta gaa aaa gca gcg tct	50	55	60	192
Asp Gln Lys Lys Thr Glu Glu Lys Ala Ala Leu Glu Lys Ala Ala Ser				
gaa gag atg gat aag gca gtg gca gca gtt caa caa gcg tat cta gcc	65	70	75	240
Glu Glu Met Asp Lys Ala Val Ala Ala Val Gln Gln Ala Tyr Leu Ala				
tat caa caa gct aca gac aaa gcc gca aaa gac gca gca gat aag atg	85	90	95	288
Tyr Gln Gln Ala Thr Asp Lys Ala Ala Lys Asp Ala Ala Asp Lys Met				
ata gat gaa gct aag aaa cgc gaa gaa gag gca aaa act aaa ttt aat	100	105	110	336
Ile Asp Glu Ala Lys Lys Arg Glu Glu Ala Lys Thr Lys Phe Asn				
act gtt cga gca atg gta gtt cct gag cca gag cag ttg gct gag act	115	120	125	384
Thr Val Arg Ala Met Val Val Pro Glu Pro Glu Gln Leu Ala Glu Thr				
aag aaa aaa tca gaa gaa gct aaa caa aaa gca cca gaa ctt act aaa	130	135	140	432
Lys Lys Lys Ser Glu Glu Ala Lys Gln Lys Ala Pro Glu Leu Thr Lys				
aaa cta gaa gaa gct aaa gca aaa tta gaa gag gct gag aaa aaa gct	145	150	155	480
Lys Leu Glu Glu Ala Lys Ala Lys Leu Glu Glu Ala Glu Lys Lys Ala				
act gaa gcc aaa caa aaa gtg gat gct gaa gaa gtc gct cct caa gct	165	170	175	528
Thr Glu Ala Lys Gln Lys Val Asp Ala Glu Glu Val Ala Pro Gln Ala				
aaa atc gct gaa ttg gaa aat caa gtt cat aga cta gaa caa gag ctc	180	185	190	576
Lys Ile Ala Glu Leu Glu Asn Gln Val His Arg Leu Glu Gln Glu Leu				
aaa gag att gat gag tct gaa tca gaa gat tat gct aaa gaa ggt ttc	195	200	205	624
Lys Glu Ile Asp Glu Ser Glu Ser Glu Asp Tyr Ala Lys Glu Gly Phe				
cgt gct cct ctt caa tct aaa ttg gat gcc aaa aaa gct aaa cta tca	210	215	220	672
Arg Ala Pro Leu Gln Ser Lys Leu Asp Ala Lys Lys Ala Lys Leu Ser				
aaa ctt gaa gag tta agt gat aag att gat gag tta gac gct gaa att	225	230	235	720
Lys Leu Glu Glu Leu Ser Asp Lys Ile Asp Glu Leu Asp Ala Glu Ile				
gca aaa ctt gaa gat caa ctt aaa gct gct gaa gaa aac aat aat gta	245	250	255	768
Ala Lys Leu Glu Asp Gln Leu Lys Ala Ala Glu Glu Asn Asn Asn Val				
gaa gac tac ttt aaa gaa ggt tta gag aaa act att gct gct aaa aaa	260	265	270	816
Glu Asp Tyr Phe Lys Glu Gly Leu Glu Lys Thr Ile Ala Ala Lys Lys				

gct gaa tta gaa aaa act gaa gct gac ctt aag aaa gca gtt aat gag	275	280	285	864
Ala Glu Leu Glu Lys Thr Glu Ala Asp Leu Lys Lys Ala Val Asn Glu				
cca gaa aaa cca gct cca gct cca gaa act cca gcc cca gaa	290	295	300	906
Pro Glu Lys Pro Ala Pro Ala Pro Glu Thr Pro Ala Pro Glu				
<210> 27				
<211> 302				
<212> PRT				
<213> <i>Streptococcus pneumoniae</i>				
<400> 27				
Glu Glu Ser Pro Val Ala Ser Gln Ser Lys Ala Glu Lys Asp Tyr Asp	1	5	10	15
Ala Ala Lys Lys Asp Ala Lys Asn Ala Lys Lys Ala Val Glu Asp Ala	20	25	30	
Gln Lys Ala Leu Asp Asp Ala Lys Ala Ala Gln Lys Lys Tyr Asp Glu	35	40	45	
Asp Gln Lys Lys Thr Glu Glu Lys Ala Ala Leu Glu Lys Ala Ala Ser	50	55	60	
Glu Glu Met Asp Lys Ala Val Ala Val Gln Gln Ala Tyr Leu Ala	65	70	75	80
Tyr Gln Gln Ala Thr Asp Lys Ala Ala Lys Asp Ala Ala Asp Lys Met	85	90	95	
Ile Asp Glu Ala Lys Lys Arg Glu Glu Ala Lys Thr Lys Phe Asn	100	105	110	
Thr Val Arg Ala Met Val Val Pro Glu Pro Glu Gln Leu Ala Glu Thr	115	120	125	
Lys Lys Lys Ser Glu Glu Ala Lys Gln Lys Ala Pro Glu Leu Thr Lys	130	135	140	
Lys Leu Glu Glu Ala Lys Ala Lys Leu Glu Glu Ala Glu Lys Lys Ala	145	150	155	160
Thr Glu Ala Lys Gln Lys Val Asp Ala Glu Glu Val Ala Pro Gln Ala				

165

170

175

Lys Ile Ala Glu Leu Glu Asn Gln Val His Arg Leu Glu Gln Glu Leu  
180 185 190

Lys Glu Ile Asp Glu Ser Glu Ser Glu Asp Tyr Ala Lys Glu Gly Phe  
195 200 205

Arg Ala Pro Leu Gln Ser Lys Leu Asp Ala Lys Lys Ala Lys Leu Ser  
210 215 220

Lys Leu Glu Glu Leu Ser Asp Lys Ile Asp Glu Leu Asp Ala Glu Ile  
225 230 235 240

Ala Lys Leu Glu Asp Gln Leu Lys Ala Ala Glu Glu Asn Asn Asn Val  
245 250 255

Glu Asp Tyr Phe Lys Glu Gly Leu Glu Lys Thr Ile Ala Ala Lys Lys  
260 265 270

Ala Glu Leu Glu Lys Thr Glu Ala Asp Leu Lys Lys Ala Val Asn Glu  
275 280 285

Pro Glu Lys Pro Ala Pro Ala Pro Glu Thr Pro Ala Pro Glu  
290 295 300